This is a pre-copyedited, author-produced version of an article accepted for publication in Journal of Economic Geography following peer review. The version of record: Roberto Ezcurra, Andrés Rodríguez-Pose, Does ethnic segregation matter for spatial inequality?, Journal of Economic Geography, Volume 17, Issue 6, November 2017, Pages 1149–1178, doi 10.1093/jeg/lbx007 is available online at:
https://doi.org/10.1093/jeg/lbx007

Does ethnic segregation matter for spatial inequality? A cross-country analysis

Roberto Ezcurra
Universidad Pública de Navarra
roberto.ezcurra@unavarra.es

Andrés Rodríguez-Pose
London School of Economics
a.rodriguez-pose@lse.ac.uk

Abstract

The paper examines the link between ethnic segregation and spatial inequality in 71 countries with different levels of economic development. The results reveal that ethnic segregation is associated with significantly higher levels of spatial inequality. This finding is not affected by the inclusion of various covariates that may influence both spatial inequality and the geographical distribution of ethnic groups, and is confirmed by a number of robustness tests. The results also suggest that political decentralization and the quality of government could act as transmission channels linking ethnic segregation and spatial inequality.

Keywords: spatial inequality, ethnic segregation.

JEL classification: J15, O11, 018, R11.

1 Introduction

Spatial inequality has recently received considerable attention from scholars and decision-makers. The growing interest surrounding this issue derives from the fact that spatial inequality, defined as income inequality across geographical or administrative units within a territorial entity (e.g. country, region), is a crucial component of overall inequality (Milanovic, 2005). People with similar income levels often concentrate in space, meaning that increases in spatial inequality are associated –other things being equal– with greater levels of interpersonal inequality. Spatial inequality is also important because high regional income disparities may lead to internal conflicts, undermining social and political stability (Kanbur and Venables, 2005; Østby et al., 2009).

The impact of different factors on spatial inequality has been profusely examined. The focus has been mainly on the level of economic development (Petrakos et al., 2005; Lessmann, 2014), the degree of trade openness (Rodríguez-Pose, 2012; Ezcurra and Rodríguez-Pose, 2014a), or fiscal and political decentralization (Rodríguez-Pose and Ezcurra, 2010; Kyriacou et al., 2015). However, the influence of the ethnic composition of the population has attracted much less attention. Only a limited number of studies include an index of ethnolinguistic fractionalization as an additional control within models which deal with different research questions (e.g. Lessmann, 2012, 2014; Ezcurra and Rodríguez-Pose, 2013). This type of indices, by contrast, has been widely used to investigate the effect of ethnic diversity on other aspects of economic and political performance, starting with the influential contribution by Easterly and Levine (1997) (see Alesina and La Ferrara (2005) for a survey of this literature). One of the problems with fractionalization indices is that, although they contain information about the identity and size of the various ethnic and/or linguistic groups, they incorporate no additional information about other substantive characteristics of the groups. In particular, these indices do not capture the extent to which the members of each group are spatially clustered, ignoring the degree of segregation within countries (Alesina and La Ferrara, 2005; Alesina and Zhuravskaya, 2011). Nevertheless, ethnic segregation is an important aspect of regional diversity and there are reasons to expect that it may affect spatial inequality, as ethnic groups often differ in economic terms (Baldwin and Huber, 2010; Alesina et al., 2016).

However, despite its importance, existing cross-country studies focusing on spatial inequality have tended to overlook ethnic segregation. To the best of our knowledge, only Kyriacou and Roca-Sagalés (2012) and Kyiriacou et al. (2015) include a measure of segregation when analysing the effect on regional disparities of the EU structural funds and fiscal decentralization. Nevertheless, the primary focus of these papers is not the study of the link between ethnic segregation and spatial inequality. Furthermore, the samples in Kyriacou and Roca-Sagalés (2012) and Kyiriacou et al. (2015) comprise respectively EU and OECD countries alone. These are countries where ethnic segregation and regional disparities are considerably lower than in the developing world (Alesina and Zhuravskaya, 2011; Ezcurra and Rodríguez-Pose, 2013). In addition to these two papers, in one of the robustness tests included in a study on the link between the quality of government and spatial inequality, Ezcurra and Rodríguez-Pose (2014b) also control for the degree of segregation. However, their analysis covers only

46, mostly well-developed, countries.

Hence, as far as we are aware, no paper has directly addressed the link between ethnic segregation and within-country spatial inequality across a large number of countries in the world. We aim to fill this gap in the literature by examining the extent to which the geographical concentration of ethnic groups shapes regional income disparities. In order to reach this goal, the paper investigates the relationship between ethnic segregation and spatial inequality in a cross-section of 71 countries with different levels of economic development. The main aim is to examine whether the geographical concentration of different ethnic groups within a country (ethnic segregation) is connected with higher levels of spatial inequality. This is particularly relevant, as the possible existence of a positive effect of segregation on regional income disparities would increase the risks that high levels of spatial inequality could be a source of social unrest and ethnic conflict (Bakke and Wibbels, 2006).

The remainder of the paper is organized as follows. After this introduction, section 2 discusses from a theoretical perspective the relationship between ethnic segregation and spatial inequality. Section 3 describes the measures of ethnic segregation and spatial inequality used in the paper. In section 4 we present the results of the empirical analysis. The robustness of our findings is checked in section 5. Section 6 explores various potential transmission channels linking ethnic segregation and spatial inequality. The final section offers the main conclusions of the paper.

2 The relationship between ethnic segregation and spatial inequality

There are various reasons to assume that ethnic segregation is related to spatial inequality. As pointed out by Alesina and Zhuravskaya (2011), the geographical concentration of ethnic groups increases the risk of secession, which poses a major threat to peace and stability within a country. Secessionist groups are a force in many parts of the world, from developing –such as China, India, Indonesia, Nigeria, or Sri Lanka– to well-developed countries –such as Belgium, Canada, Spain, or the United Kingdom. The threat of secession is generally more serious if the ethnic group aiming to form its own state is spatially segregated and lives near the border of the country (Sorens, 2005; Alesina and Zhuravskaya, 2011), as in the cases of Chechnya in Russia, Xinjiang and Tibet in China, or Catalonia and the Basque Country in Spain. Faced with the risk of secession, national governments may opt for repressive policies, often leading to internal conflict with potentially important consequences in political and economic terms. An alternative option, used by numerous countries over the last decades, is to resort to fiscal redistribution and decentralization as a way to buy-back the loyalty of separatist regions (Bakke and Wibbels, 2006; Kyriacou and Morral-Palacín, 2015). Fiscal and political decentralization can help to reduce ethnic conflicts and the threat of secession by bringing the government closer to the citizens, increasing the opportunities to participate in government and giving groups greater autonomy and control over their political, social and economic affairs (Brancati, 2006).

From a political perspective, decentralization can be an useful instrument to tackle the risk of secession in countries with high levels of ethnic segregation. In addition, one may expect more segregated countries to be more decentralized for another reason. The so-called ‘decentralization theorem’ states that subnational tiers of government are more capable than central government to tailor the provision of public goods to the needs of the local population due to the existence of informational advantages and a better insight into the preferences of citizens (Tiebout, 1956; Oates, 1972). In a country where ethnic groups are geographically clustered, it is likely that the needs and preferences for public goods provision differ across regions. In such a setting, the devolution of fiscal and political power from central to subnational governments can lead to efficiency gains in the allocation of resources and in government activities (Rodríguez-Pose et al., 2009).

Decentralization may also affect regional disparities. In fact, according to the prevailing opinion, the distribution of any potential economic dividend linked to decentralization is likely to be spatially uneven (Rodríguez-Pose and Gill, 2005; Rodríguez-Pose and Ezcurra, 2010). The transfer of authority and resources to subnational tiers of government often benefits the most prosperous regions, which generally enjoy better socio-economic endowments and higher quality institutions. Decentralization processes are also likely to undermine the capacity of central governments to play an equalizing role, further increasing spatial inequality (Prud’homme, 1995). Other scholars, however, argue that decentralization may be associated with a reduction in regional income disparities, since it may contribute to a more balanced distribution of resources across space. For example, in the framework of the second generation models of fiscal federalism, Weingast (1995) and Qian and Weingast (1997) underline that fiscal decentralization fuels regional competition. Given that the ability of regional governments to stay in power depends decisively on their performance in attaining a level of development and economic growth similar to that of the rest of the country, policy-makers in poorer regions may attempt to reduce their development gaps by offering more flexible labour markets and/or less generous welfare provisions than richer regions. Likewise, decentralized countries may do better at reducing spatial inequality because of the greater political risks that regional disparities pose for such countries (Shankar and Shah, 2003).

The results of studies testing the link between decentralization and spatial inequality remain, however, inconclusive. Most analyses dealing with the developed world find that decentralization contributes to regional convergence (e.g. Ezcurra and Pascual, 2008; Lessmann, 2012). Rodríguez-Pose and Gill (2004), in a study that covers developed and developing countries, find mixed evidence depending on the countries, but a prevailing tendency for the devolution of political and fiscal powers to be associated with increasing regional disparities. In turn, the results in Rodríguez-Pose and Ezcurra (2010) and Kyriacou et al. (2015) suggest that the relationship between the processes of decentralization and spatial inequality may be contingent on the level of development and the quality of government.

Furthermore, Alesina and Zhuravskaya (2011) argue that segregation adversely affects government quality. In particular, these authors find that higher segregation in terms of ethnicity and language is associated with significantly lower quality of government. A first channel that could explain the link between segregation and governance is trust. There is abundant evidence showing that individuals prefer to interact and associate with other members of their own ethnic group (Glaeser et al., 2000; Costa and Kahn, 2003). Consequently, countries where ethnic groups are geographically clustered tend to register less social interactions between members of different groups, which leads to lower levels of trust –a key component of social capital– in the country as a whole. Indeed, Uslaner (2008) and Alesina and Zhuravskaya (2011) have confirmed empirically the presence of a negative correlation between segregation and generalized trust. In turn, numerous studies show a negative impact of low trust on the quality of government (e.g. Putnam, 1993; La Porta et al. 1997).

In addition to trust, Alesina and Zhuravskaya (2011) identify other factors that may explain the observed link between segregation and quality of government. As discussed above, the geographical concentration of ethnic groups could increase the development of independence movements and the threats of secession. In this setting, national governments may have incentives to spend additional resources in repressive measures or interregional redistribution initiatives, crowding out other policies that contribute to improving the overall quality of government (Alesina and Spolaore, 1997, 2003; Kyriacou and Morral-Palacín, 2015). Moreover, segregation tends to increase the relevance of ethnic voting, which may harm institutional quality, especially in democratic regimes where rules are chosen through competitive elections (Chandra, 2004; Banerjee and Pande, 2007).

A negative association between segregation and the quality of government is particularly important for our study, as government quality may in turn be related to spatial inequality (Ezcurra and Rodríguez-Pose, 2014b; Kyriacou et al., 2015). Countries with weak institutions and low quality of government are characterized by the presence of persistent corruption, pervasive rent-seeking, self-serving decision-makers, and low quality of bureaucracy. This set of problems often gives rise to imperfectly functioning markets and institutional and government failure, which reduces the capacity of the public sector to design and implement effective policies that contribute to improving regional convergence and reducing spatial inequality. Government quality also plays an essential role in establishing the adequate conditions for economic interactions and reducing the risks of social unrest and political instability. By decreasing uncertainty and transaction costs, governments can facilitate the processes of technology and knowledge transfer across regions, improving the conditions for the development of economic activity in lagging regions (North, 1990; Acemoglu and Robinson, 2012).

The previous discussion shows that ethnic segregation may affect spatial inequality through its impact on the processes of decentralization and the quality of government. However, the different arguments laid down above also show that this is a complex relationship involving multiple factors and mechanisms, which are often interdependent. Empirical research is therefore needed in order to shed light on this issue. For this reason, the rest of the paper is devoted to investigating the link between ethnic segregation and spatial inequality in a cross-section of countries.

3 Measuring ethnic segregation and spatial inequality

To measure segregation, we resort to the index proposed by Alesina and Zhuravskaya (2011):

$$S_i = \frac{1}{M-1} \sum_{m=1}^{M} \sum_{j=1}^{J} \frac{p_j(\pi_{jm} - \pi_m)^2}{\pi_m} \tag{1}$$

where p_j is the population share of region j in country i, π_m is the fraction of group m in country i, and π_{jm} is the fraction of group m in region j of country i. M and J represent respectively the total number of groups and regions in country i. Crucially for the purpose of this paper, S includes a spatial dimension, as it allows to quantify the degree of geographical concentration of the different ethnic, religious, or linguistic groups within a country. The value of the index ranges from zero when every region has the same share of each group as the country as a whole (no segregation), to one when each region is inhabited by a separate group (full segregation).

When calculating S, it has to be borne in mind that in many regions of different countries a fraction of the population remains not ascribed to any particular group, generally under the 'other' category. We assume that the 'other' group is comprised of a number of distinct and small subgroups O which cannot be classified adequately due to the lack of data. We also assume that the different subgroups included in the 'other' category are uniformly distributed across all regions in the country, implying no segregation within the 'other' category. Under these two assumptions, the segregation index S can be rewritten as follows (Alesina and Zhuravskaya, 2011):

$$\hat{S}_i = \frac{1}{N+O-1} \left(\sum_{m=1}^{M} \sum_{j=1}^{J} \frac{p_j(\pi_{jm} - \pi_m)^2}{\pi_m} + S_o \right) \tag{2}$$

with

$$S_o = \sum_{j=1}^{J} \frac{p_j(\pi_{jo} - \pi_o)^2}{\pi_o} \tag{3}$$

where N is the number of identified groups, π_{jo} is the fraction of 'others' in region j, and π_o is the fraction of 'others' in the whole population. In the rest of the paper we focus on $\hat{S}$ as our main measure of segregation.

Applying the same classification of groups used in Alesina et al. (2003), Alesina and Zhuravskaya (2011) measure the level of segregation for three different dimensions of diversity: ethnicity, language, and religion. In this paper we use their indices of ethnic segregation, which combine language, self-reported ethnicity, and physical features (primarily skin colour). To calculate the indices, Alesina and Zhuravskaya (2011) collect data for subnational administrative units within each country (i.e. regions) on the total population size and the share of the population that belongs to the various groups. These data are drawn from the census closest to the year 2000, whenever it was available, national statistical offices, and demographic and health surveys. Using these sources, Alesina and Zhuravskaya (2011) compute their measure of ethnic segregation for 97 countries.

In order to measure regional income disparities within each country, we resort to the population-weighted coefficient of variation. This measure of dispersion, which is widely used in the literature on regional disparities (e.g. Williamson, 1965; Rodríguez-Pose and Ezcurra, 2010; Lessmann, 2014), adopts the following form:

$$c_i = \frac{1}{\mu}\left[\sum_{j=1}^{J} p_j \left(y_j - \mu\right)^2\right]^{0.5} \tag{4}$$

where y and p are respectively the GDP per capita and population share of region j in country i, and $\mu = \sum_{j=1}^{J} p_j y_j$. The advantage of this inequality measure vis-à-vis other potential alternative measures is that it is independent of scale and population size, and satisfies the Pigou-Dalton transfer principle (Cowell, 1995). Moreover, it is not sensitive to the number of regions within each country (Portnov and Felsenstein, 2005) and takes into account differences in population size across the regions considered. This latter aspect has traditionally been overlooked by the literature on economic convergence, despite the fact that omitting population size significantly biases our perceptions of spatial inequality (see Petrakos et al., 2005; Lessmann, 2014).

In order to calculate c and match it with Alesina and Zhuravskaya's (2011) segregation index, we require regional data on GDP and population involving the same level of territorial disaggregation used to compute $\hat{S}$. This is problematic, because although the OECD or Eurostat provide regional data for the majority of developed countries, lack of adequate regional data in developing countries represents a serious barrier to the analysis. Faced with this problem of information availability, we resort to the regional data collected by Ezcurra and Rodríguez-Pose (2013) and Gennaioli et al. (2013). Using these datasets, we calculate the c spatial inequality index for 71 of the 97 countries in Alesina and Zhuravskaya's (2011) sample (see Table A1 in the Appendix for further details). The level of territorial disaggregation considered in the analysis usually coincides with the highest administrative division available (i.e., states and provinces rather than counties or municipalities) and in a few cases with the statistical division (e.g., the Eurostat NUTS in Europe) that is closest to it.[1] When

[1]When interpreting the results of the paper, it is important to note that our findings depend ultimately on the level of territorial disaggregation used to measure both ethnic segregation and spatial inequality. This is relevant because in some countries ethnic segregation and spatial inequality may occur at a much finer level of territorial disaggregation than that considered in our study, which cannot be adequately captured by the dataset used in the paper.

possible, we take the mean value of c over the period 2000-2005 as a baseline, in order to reduce the potential bias due to the impact of the business cycle on regional income disparities (Lessmann, 2014).

Tables 1 and 2 show the countries with the highest and the lowest values of $\hat{S}$ and c. Ethnic segregation ranges from 0.001 (Germany) to 0.394 (Zimbabwe) with a mean value of 0.088 and a standard deviation of 0.105, whereas spatial inequality ranges from 0.076 (Australia) to 1.112 (Ecuador) with a mean value of 0.375 and a standard deviation of 0.202. Both ethnic segregation and spatial inequality tend to be considerably higher in developing countries than in developed ones. Of the ten countries with the highest levels of ethnic segregation in the sample, only Spain –a country with sizeable separatist movements– is high income. In turn, Paraguay is the only developing country with one of the ten least segregated populations in the sample. Similarly, the states with the highest levels of spatial inequality are all developing countries, with the exceptions of Russia and Latvia. Malawi, Senegal, and Armenia are the only developing countries among those with low levels of spatial inequality.

[INSERT TABLE 1 AROUND HERE]

[INSERT TABLE 2 AROUND HERE]

Given that the aim of the paper is to assess the potential link between ethnic segregation and spatial inequality, we divide the sample countries into two and three groups according to the value of $\hat{S}$ in order to get a preliminary insight into this relationship. The definitions of the different groups are based on the median (classification into two groups) and the first and third quartiles (classification into three groups) of the distribution of the index of segregation. The results in Table A2 in the Appendix reveal that more segregated countries have, on average, higher levels of spatial inequality. By contrast, low ethnic segregation is associated with lower regional income disparities. This is corroborated by the corresponding F-tests, which show that the differences between the groups in the mean value of c are statistically significant at the 1% level. This preliminary evidence suggests a positive correlation between ethnic segregation and spatial inequality.

The descriptive results of Table A2 may, however, be sensitive to the specific number of groups used to classify the sample countries. More importantly, they may be biased by the fact that spatial inequality is not only dependent on the degree of ethnic segregation. This analysis thus provides only a preliminary view of the potential direction of the relationship between ethnic segregation and spatial inequality, as omitted variables can seriously affect our perception of how the geographical distribution of ethnic groups shapes regional disparities. In the remainder of the paper we carry out a more appropriate statistical analysis.

4 Is there a link between ethnic segregation and spatial inequality? Empirical analysis

4.1 The model

Studying the relationship between ethnic segregation and spatial inequality, leads us to estimate different versions of the following model:

$$\log I_i = \alpha + \beta S_i + \gamma^{'} \mathbf{X}_i + \varepsilon_i \tag{5}$$

where I is the measure of spatial inequality in country i, S is the index of ethnic segregation, $\mathbf{X}$ denotes a set of variables controlling for additional factors assumed to influence regional income disparities, including continent dummies for Latin America, North America, Europe, Sub-Saharan Africa, Asia, and Pacific. ε is the corresponding disturbance term. The coefficient of interest throughout the paper is β, which measures the effect of ethnic segregation on spatial inequality.

The control variables in vector $\mathbf{X}$ have been selected on the basis of existing studies on the determinants of spatial inequality. They include the average size of the regions used in each country to compute the magnitude of spatial inequality, the degree of ethnic fractionalization of the population, the stage of economic development, trade openness, country size, and a dummy variable for transition countries. The definitions of all the control variables and their sources are presented in the Appendix.

When estimating model (5), the level of regional income disparities in each country may be affected by the average size of the territorial units used to compute the spatial inequality index. This is especially relevant in our analysis, as the average size of the territorial units used to calculate c differs considerably among countries. In addition, the size of a country's regions may have a direct effect on the measure of ethnic segregation (Alesina and Zhuravskaya, 2011). Hence and despite the fact that the values of the dependent variable have already been calculated taking into account the differences in population size across regions, we also control for the average size of regions in each country as a way to minimize any potential bias emerging from the heterogeneity in territorial levels across countries.

According to Lessmann (2012, 2014) and Ezcurra and Rodríguez-Pose (2013), the degree of ethnic fractionalization of the population may also affect the level of spatial inequality within any given country. The findings of these authors seem to indicate that those countries with a higher degree of fractionalization have greater regional income disparities. As discussed in the introduction, segregation and fractionalization are two different notions (Alesina and La Ferrara, 2005). Nevertheless, Alesina and Zhuravskaya (2011) show that they are positively correlated. We therefore control for the degree of ethnic fractionalization in the sample countries by means of the index of ethnic fractionalization compiled by Alesina and Zhuravskaya (2011). These authors employ the regional data used to compute the measure of segregation described above to construct an index of fractionalization at the national level for each country. This index measures the probability that two randomly selected individuals in a given country belong to different ethnic groups.[2]

Since the pioneering work of Williamson (1965), analyses of spatial inequality have taken into consideration the level of economic development (Lessmann, 2014). Factors such as the existence of diseconomies of agglomeration, core-periphery spread effects, technological diffusion processes, or transport infrastructures affecting the location of private capital suggest that, beyond a threshold level, advances in the economic development process may contribute to the spatial dispersion of economic activity (Thisse, 2000; Petrakos et al., 2005). Other approaches, including the endogenous growth school and the so-called new economic geography, by contrast, highlight that economic growth is often associated with uneven spatial development (Krugman, 1998; Fujita and Thisse, 2002). As is common practice in the literature, we use GDP per capita to capture existing differences in development across the sample countries. Moreover, following Lessmann (2014), we include in the list of regressors of model (5) the share of population living in urban areas as a proxy for potential agglomeration effects, which may affect regional income disparities. Countries with higher urbanization rates tend also to be characterized by lower levels of ethnic segregation, as group mixing is more likely in the cities (Alesina and Zhuravskaya, 2011).

The rise in trade over recent decades has also attracted considerable attention as a potential determinant of regional income disparities (Kanbur and Venables, 2005; Rodríguez-Pose, 2012). The impact of trade on spatial inequality is, however, not entirely well understood. On the one hand, Heckscher-Ohlin type analyses indicate that trade contributes to reducing existing disparities, in the cases when capital investment is attracted by those regions with the lowest cost base, while labour moves to the highest salary zones. On the other hand, according to this theory, the owners of abundant factors will benefit from trade, but owners of scarce resources will experience falling returns, at least in the medium term. New economic geography models provide different outcomes in relation to the link between trade and spatial inequality, depending on the theoretical assumptions employed in each case. We therefore control for the possible impact on spatial inequality of the degree of trade openness, measured as the ratio between total trade (exports and imports) and GDP.

Spatial inequality may also be related to country size (Williamson, 1965). Larger countries are often characterized by greater spatial heterogeneity than smaller countries, which are in general more homogeneous and compact. We use the area of a country as our measure of country size.

Transition from real socialism to capitalism is also bound to have affected the location of economic activities and thus spatial inequality. Throughout the 1990s, a number of countries around the world –and especially in Central and Eastern Europe– underwent profound changes of a political and economic nature as a consequence of

[2]The index of fractionalization can be expressed as follows:

$$F_i = \sum_{m=1}^{M} \pi_m (1 - \pi_m)$$

the processes of restructuring, privatization, and liberalization that ensued the fall of communism. These changes have had a significant impact on the spatial distribution of economic activity, frequently leading to an important increase in the magnitude of regional income disparities (Ezcurra and Pascual, 2007). Consequently, a dummy variable for transition countries is included in vector X.

4.2 Results

Table 3 presents the results of estimating various versions of model (5) using OLS with heteroskedasticity robust standard errors. The coefficient of the index of ethnic segregation, $\hat{S}$, is in all cases positive and statistically significant at the 1% level. This implies that more ethnically segregated countries have on average higher levels of spatial inequality, corroborating the preliminary evidence of Table A2. The R-squared in column (1) of Table 3 indicates that $\hat{S}$ alone explains around 16% of the variation in the dependent variable. At this point it is important to recall that Alesina and Zhuravskaya (2011) show that both the degree of ethnic fractionalization and the level of economic development are correlated with segregation. Nevertheless, Table 3 reveals that $\hat{S}$ remains significantly associated with spatial inequality, even when we control for GDP per capita and fractionalization.[3] Hence, ethnic segregation contributes to explaining the variation in spatial inequality and does not simply capture the effect of these two covariates. The inclusion of other explanatory variables in the analysis also does not alter the observed relationship, confirming its robustness and showing that the effect of ethnic segregation on spatial inequality is not a spurious correlation resulting from the omission of relevant variables. Figure A1 in the Appendix illustrates the link between ethnic segregation and regional disparities with a partial regression plot based on all covariates.

[INSERT TABLE 3 AROUND HERE]

Regarding the control variables in model (5), Table 3 reveals that our results are in general consistent with the findings in the literature on the determinants of spatial inequality. There is a negative association between the average size of regions and the level of spatial inequality in a country. Our estimates also show that more ethnically homogeneous countries tend to experience lower levels of regional income disparities, whereas the coefficients of GDP per capita and the share of urban population are not statistically significant consistently across the various regressions.[4] Furthermore, the relationship between trade openness and spatial inequality is always positive and statistically significant, confirming the empirical evidence provided by Rodríguez-Pose

[3]The inclusion of GDP per capita in model (5) is, however, controversial, as this variable may be a proximate outcome of segregation and fractionalization (Angrist and Pischke, 2009; Alesina and Zhuravskaya, 2011).

[4]We investigate the possibility that the effect of GDP per capita on regional income disparities may be non-linear (Lessmann, 2014). The results show that the coefficient of the index of ethnic segregation remains positive and statistically significant, but the estimates do not support the hypothesis of a non-linear link between GDP per capita and spatial inequality in our sample.

and Ezcurra (2010) and Ezcurra and Rodríguez-Pose (2013). The magnitude of regional income disparities is also greater in larger countries and transition economies.

Overall, the results in Table 3 show a strong positive correlation between ethnic segregation and spatial inequality. Nevertheless, there are important reasons that prevent us from interpreting this relationship as causal at this stage. First, the segregation observed in a country is contingent on where people live, and this choice may be endogenous to the spatial distribution of income. Consequently, segregation may affect regional disparities and, in turn, be affected by them, triggering a reverse causality problem. Moreover, the measures of ethnic segregation and spatial inequality are conditional on internal administrative boundaries. This implies that there may be omitted determinants of spatial inequality that will naturally be correlated with segregation. Furthermore, the index of ethnic segregation may be affected by measurement error, which may bias the OLS estimates downward. All of these problems are potentially important from an econometric perspective, but can be solved with an appropriate instrument for segregation. Such an instrument must not be correlated with the disturbance term in model (5), but account for the variation in segregation observed in the sample. Fortunately, we can use the instrument for ethnic segregation proposed by Alesina and Zhuravskaya (2011). This instrument relates the spatial distribution of ethnic groups in a country to the ethnic composition of neighbouring countries. In particular, Alesina and Zhuravskaya (2011) assume that when a specific ethnic group in the home country is also present in one of the neighbouring countries, it is likely that the members of this group will tend to concentrate near the border with this neighbouring country. On the contrary, if a group in the home country is not present in any of the neighbouring states, it is likely that the members of this group will be more uniformly distributed across the country, and not located closer to any particular border. This idea is used by Alesina and Zhuravskaya (2011) to construct a predicted distribution of the different ethnic groups within the various countries, assuming that the members of a specific group 'gravitate' towards the borders of neighbouring countries that are populated by people from the same ethnic group. Using this predicted distributions, Alesina and Zhuravskaya (2011) construct an index of predicted segregation, which can be used as an instrument for actual segregation.[5]

In order to assess whether the instrument is correlated with ethnic segregation, Table 4 presents the results of the first stage regressions of the form:

$$S_i = \delta + \zeta S_i^p + \theta' \mathbf{X}_i + \upsilon_i \tag{6}$$

where S_i^p is the predicted segregation index and υ is the corresponding error term. The table is organized in the same way as Table 3. As can be observed, in all the regressions the instrument has a positive and statistically significant effect on actual segregation. This relationship is illustrated in Figure A2 in the Appendix, which plots predicted versus actual segregation, conditional on the full set of control variables. Table 4

[5]See Alesina and Zhuravskaya (2011, pp.1889-1893) for further details and examples on the construction of the instrument.

also reports the F-statistics for the excluded instrument calculated both under the assumptions of homoskedasticity and heteroskedasticity in the error term. Although the former is an incorrect assumption in this case, we include it because the cutoff points proposed by the theory on weak instruments are based only on the homoskedasticity assumption (Stock et al., 2002). Under this assumption, the F-statistics of the first stage regressions are in all cases cases well above the threshold of 10 suggested by Staiger and Stock (1997) when there is a single endogenous regressor. These results are corroborated by the partial R-squared, which measures the correlation between $\hat{S}$ and the instrument after partialling out the effect of the remaining regressors.[6]

[INSERT TABLE 4 AROUND HERE]

The information in Table 4 indicates that the index of predicted segregation is significantly associated with actual segregation. To be a valid instrument, however, the index of predicted segregation should not affect spatial inequality, beyond its impact through actual segregation. This condition cannot be tested formally in the absence of other instruments. We therefore discuss various possible arguments that may cast doubt on the excludability of the instrument. First, national borders do not always take into account the historical location of ethnic groups, meaning that boundary lines can split ethnic groups into different countries (Alesina et al., 2011). Although there are other examples of this type of artificial borders around the world, this situation is particularly notorious in Sub-Saharan Africa (Herbst, 2000; Englebert et al., 2002). In most of Sub-Saharan African countries, the borders were drawn during the colonial period, following in many cases parallels and meridians, and few borders changed after decolonization. In this setting, the presence of the same ethnic group in two adjacent countries may influence the index of predicted segregation used as instrument and affect spatial inequality in other ways than through ethnic segregation. For example, one may expect that inter-country mobility and trade are more likely when the same ethnic group inhabits on both sides of the border. In order to address empirically this concern, we use different strategies. First, we repeat the analysis excluding Sub-Saharan African countries, where arbitrary borders are especially relevant. Second, we add to our baseline specification various measures of artificial borders, including the share of a country's population that belongs to ethnic groups that are also present in neighbouring countries. Third, taking into account the possibility that the ethnic composition of neighbouring countries may influence spatial inequality in the home country through spatial spillovers across the national borders, we also control for the level of segregation and fractionalization in neighbouring countries. As can be seen in section 5, the observed relationship between ethnic segregation and spatial inequality survives these tests.

Furhtermore, the instrument could also be correlated with the length of state history and with the process of state formation. In particular, given the nature of the

[6] In order to confirm that the observed relationship between ethnic segregation and spatial inequality is not affected by the employment of a potentially weak instrument, we have checked that our results still hold when we calculate for all the IV regressions presented in the paper the corresponding confidence intervals using the method proposed by Finlay and Magnusson (2009). This method provides reliable confidence intervals in the case of weak instruments and is robust to arbitrary heteroskedasticity.

instrument, one may expect it to be more relevant in countries with a shorter state history, which tend to be characterized by higher levels of ethnic diversity (Ahlerup and Olsson, 2012), as ethnic identification can be considered a consequence of modern state formation (Gellner, 1983; Hobsbawm and Ranger, 1983). This has been the case of countries such as France, Spain or Great Britain, where the statehood experience contributed to national identity formation by forging state-level institutions and adopting active policies designed to reduce ethnic heterogeneity and bring about more homogeneous populations (Tilly, 1992). This process of state formation contrasts with the experience of most of the European colonies, which were created and administered with the main purpose of benefiting the colonial powers. The Europeans in charge of the colonial states tended in many cases to use the 'divide-and-rule' principle, exploiting existing ethnic conflicts to keep colonies under control, as incentives to ethnically homogenize colonial territories were limited (Ahlerup and Olsson, 2012). These considerations are potentially important in our context. Acemoglu et al. (2001, 2002), Bockstette et al. (2002) or Chanda and Putterman (2007), among others, show that state history and European colonization are related to a country's economic outcomes, which in turn may affect its level of spatial inequality. We deal with this potential problem by including in the analysis the length of state history and the colonial experience of the various countries. As described in detail in section 5, our main results still hold when we control for these factors.

Table 5 presents the results of the second stage regressions. As in the OLS regressions reported in Table 3, the coefficient of the index of ethnic segregation is in all cases positive and statistically significant. In fact, its size is in general somewhat larger than in the OLS estimates reported in Table 3. This may have to do with a potential attenuation bias generated by measurement error in the index of ethnic segregation. The regression coefficient from our preferred specification in Table 5 (column 6) reveals that raising the segregation index by one standard deviation is associated with an increase in the level of spatial inequality of around 19%. To get a more accurate idea of the dimension of the impact of ethnic segregation on regional income disparities, we consider the case of Ghana. Ghana is a country characterized by a medium level of spatial inequality ($c = 0.385$), whereas the value of the index of ethnic segregation in that country is above the sample mean ($\hat{S} = 0.112$). Our estimates indicate that if Ghana had an index of ethnic segregation similar to that registered for instance by Cameroon ($\hat{S} = 0.042$), its degree of spatial inequality could be reduced by around 13%. These figures suggest that ethnic segregation has a relevant impact on existing levels of spatial inequality.[7]

[7] As mentioned in the introduction, our analysis focuses exclusively on the relationship between ethnic segregation and spatial inequality. Alesina and Zhuravskaya (2011) also calculate different indices of linguistic and religious segregation. We have repeated the previous analysis using these alternative measures. According to the information provided by Table A3 in the Appendix, the OLS regressions show a positive and statistically significant correlation between linguistic segregation and spatial inequality. However, in the 2SLS regressions the coefficients of the indices of linguistic and religious segregation are not significant. This suggests that ethnicity is the most relevant dimension of diversity in the relationship between segregation and spatial inequality. However, this conclusion should be treated with some caution, as the identification of people with a particular ethnic group is frequently based on sharing a common mother tongue.

[INSERT TABLE 5 AROUND HERE]

5 Robustness checks

So far the analysis has revealed a positive and statistically significant effect of ethnic segregation on spatial inequality. In this section we test the robustness of this finding.

5.1 Influential observations and impact of different groups of countries

As a first robustness test we examine the impact of influential observations on the above results. We check that our findings are robust to the exclusion of any particular country from the sample. The most influential observations in 'favour' of our results are Belgium and Ecuador, two countries characterized by high levels of ethnic segregation and spatial inequality conditional on the other covariates. In turn, the most influential observations 'against' the results are Turkey and Zimbabwe, two countries with high ethnic segregation but an intermediate level of spatial inequality conditional on the remaining covariates. Although the quantitative impact of ethnic segregation on spatial inequality is affected when these countries are excluded from the sample, the observed relationship still holds. In order to confirm these findings, we resort to quantile regression as an alternative to identify the possible influence of potential outliers in our sample. Table A4 in the Appendix shows the results obtained when this method is used to estimate model (5). As can be checked, the coefficient of the index of ethnic segregation is positive and statistically significant for the various conditional quantiles of the dependent variable, which confirms that the basic nature of our findings is unaffected by the presence of potential outliers. That said, the estimates in Table A4 also indicate that the quantitative impact of ethnic segregation appears to be higher in the upper quantiles of the conditional distribution of spatial inequality.

As an additional sensitivity check, we assess the extent to which our results are determined by the inclusion of specific groups of countries in the sample. The positive association detected between ethnic segregation and spatial inequality may be driven by a particular group of countries. If this hypothesis holds, eliminating that group of countries would render the coefficient of the ethnic segregation index non-significant. We therefore re-estimate our baseline specification excluding different groups of countries (Latin America, North America, Europe, Sub-Saharan Africa, Asia, Pacific, and Middle East and North Africa). Despite the reduction of the sample, Table 6 show that the coefficient of the index of ethnic segregation remains positive and statistically significant in all cases. Taking into account the previous discussion on the excludability of the instrument, it is important to note that the results still hold when Sub-Saharan African countries, where the presence of artificial borders is especially relevant, are excluded.

[INSERT TABLE 6 AROUND HERE]

5.2 Alternative measures of spatial inequality and ethnic segregation

The results of the analysis may be also sensitive to the choice of the measure used to quantify spatial inequality. Different inequality measures may yield different orderings of the distributions, as each index has a different way of aggregating the information contained in the distribution (Cowell, 1995). We therefore calculate for each country the Theil's first measure of inequality ($T(1)$), and the standard deviation of the logarithm of regional GDP per capita (s).[8] The results of estimating model (5) using successively $T(1)$ and s, instead of c, as dependent variables are provided in Table A5 in the Appendix. The coefficient of $\hat{S}$ remains positive and statistically significant, indicating that the observed connection between ethnic segregation and spatial inequality holds regardless of the specific measure employed to calculate spatial inequality.

Our findings may also be affected by the measure of ethnic segregation used. In order to check whether that is the case, we repeat the analysis using an alternative index of segregation proposed by Alesina and Zhuravskaya (2011). This index implies ignoring the group 'other' and computing the index of segregation exclusively for the N identified groups.[9] The results in Table A5 in the Appendix indicate that the segregation–spatial inequality relationship is unaffected by the change in the index of ethnic segregation.

---

[8]These measures of inequality can be expressed as follows:

$$T(1)_i = \sum_{j=1}^{J} p_j \left(\frac{y_j}{\mu}\right) \log\left(\frac{y_j}{\mu}\right)$$

and

$$s_i = \left[\sum_{j=1}^{J} p_j \left(\log y_j - \overline{\mu}\right)^2\right]^{0.5}$$

where $\overline{\mu} = \sum_{j=1}^{J} p_j \log y_j$. $T(1)$ is a popular measure in the literature on personal income distribution and has also been used by researchers working on spatial disparities (e.g. Ezcurra and Rodríguez-Pose, 2013), while in its non-weighted version s has been widely employed in convergence analyses to capture the concept of sigma convergence (Barro and Sala-i-Martin, 1995). As in the case with the population-weighted coefficient of variation, these two indices are independent of scale and population size. $T(1)$ also fulfils the Pigou-Dalton transfer principle for the whole definition domain of income (Cowell, 1995).

[9]The resulting measure of ethnic segregation can be expressed as follows:

$$\tilde{S}_i = \frac{1}{N-1} \sum_{m=1}^{N} \sum_{j=1}^{J} \frac{p_j (\pi_{jm} - \pi_m)^2}{\pi_m}$$

When calculating $\tilde{S}$ we are omitting the fraction of the population included in the 'other' category. Accordingly, $\tilde{S}$ is only an approximation of the theoretically correct definition of segregation.

5.3 Additional controls

As an additional sensitivity check, we investigate the possibility that our results are driven by an omitted variable. This requires controlling for different covariates that may be correlated with spatial inequality, ethnic segregation, and/or the instrument used in the analysis, and checking whether the inclusion of these covariates affects our estimates.

According to the previous discussion on the excludability of the instrument of predicted segregation, we begin by examining to what extent our results are affected by the existence of artificial borders between countries that ignore the historical location of ethnic groups. Two measures of artificial states proposed by Alesina et al. (2011) are used: i) a measure quantifying the straightness (squiggliness) of borders, based on the assumption that straight borders are more likely to be artificial; ii) a measure of the share of a country's population belonging to ethnic groups partitioned by a political border. We additionally control for the length of state history using the antiquity index compiled by Putterman (2007). Taking into account that the probability of natural state formation is greater in countries that have never been colonized, we also include in the list of regressors of model (5) a dummy variable to identify former European colonies. Moreover, we examine whether the impact of the colonization process could depend on the identity of the colonizer.[10] The number of ethnic groups in a country and the ethnic composition of neighbouring states are also taken into account, as these factors may be eventually related to ethnic segregation, the instrument and spatial inequality.

Furthermore, several geographical variables are added to the baseline specification: an index of terrain ruggedness; the share of the country's surface covered by mountains, rivers and other inland bodies of water; dummy variables to identify countries with non-contiguous territories, landlocked countries, and countries surrounded by water (the instrument used in the paper predicts zero segregation for island countries); distance of the coast; and latitude. Ethnic segregation and spatial inequality may depend on the existence of physical constraints to mobility, as more topographically-uneven countries display a greater geographical concentration of economic activity (Ramcharan, 2009). Latitude is related to climatic variability and habitat diversity within a country (Cashdan, 2001), which may also affect the spatial distribution of population and income (Ahlerup and Olsson, 2012). Additionally, we include in the list of regressors the number of neighbouring countries, as the number of neighbouring countries determine the number of hypothetical regions used to calculate the index of predicted segregation used as instrument.

According to the evidence provided by Alesina et al. (2016), within-country differences in income across ethnic groups (i.e. ethnic inequality) is positively correlated with both ethnic segregation and spatial inequality. This raises the possibility that the measure of ethnic segregation used in our analysis may simply be capturing the

[10]The legal origin of the various countries was also included in order to consider any potential role of legal codes and institutions that may not be fully captured by the measures of colonial experience. However, the legal origin does not exert a statistically significant impact on spatial inequality and, most importantly, its inclusion in our baseline model does not affect the observed relationship between ethnic segregation and spatial inequality.

effect of between-ethnicity inequality. In order to explore if this is the case, we add to the baseline specification two measures of ethnic inequality calculated by Alesina et al. (2016). We also investigate whether the observed association between ethnic segregation and spatial inequality still holds when we control for the degree of ethnic tension within the various countries.

Moreover, Matuszeski and Schneider (2006) find a positive association between the incidence of civil conflicts and segregation. Given that political instability and internal conflict may affect the spatial distribution of income within a country (Abadie and Gardeazabal, 2003), we also include in the baseline model a dummy indicating if the country experienced a civil conflict during the period 1995-2005. Finally, the presence of more or less developed redistributive policies is likely to be linked to segregation and spatial inequality (Alesina and La Ferrara, 2005; Rodríguez-Pose and Ezcurra, 2010). Accordingly, any observed connection between ethnic segregation and spatial inequality may be spurious if existing differences in the capacity of the state to redistribute financial resources across regions are ignored. We therefore control for public investment and public consumption as proxies for the redistributive capacity of the government.

[INSERT TABLE 7 AROUND HERE]

The results of estimating model (5) including these additional controls are presented in Table 7. The vast majority of the newly-added controls are not statistically significant and, more importantly, have little impact on the main result of the paper. The new controls do not affect the estimates of the impact of ethnic segregation on spatial inequality. The segregation index remains positive and statistically significant in all cases, confirming, once again, the robustness of our findings.

6 Transmission channels

As described in section 2, ethnic segregation may contribute to increasing spatial inequality through its impact on the processes of decentralization, and the quality of government. In order to complement our previous findings, we now present a preliminary study on the empirical relevance of these potential transmission channels.

This analysis requires comparable data on the level of decentralization across countries. The devolution of power and responsibilities from central to regional and local governments is however a complex and multidimensional process and no single indicator captures the real level of decentralization of a country (Rodríguez-Pose and Ezcurra, 2010). We therefore resort to various indicators including standard measures of fiscal and political decentralization. In particular, our indicator of fiscal decentralization is drawn from Schneider (2003) and is based mainly on the subnational share in total government expenditure and the subnational share in total government revenue (e.g. Fisman and Gatti, 2002; Enikolopov and Zhuravskaya, 2007). Despite the popularity of these measures in the literature, they are not exempt from criticisms, as they are based exclusively on the distribution of expenditure and revenue responsibilities between the different tiers of government, but provide no information about the degree of autonomy of subnational governments (Rodríguez-Pose and Ezcurra, 2010). In order to get a more comprehensive picture of the actual powers of subnational governments, we introduce in the analysis an indicator of political decentralization proposed by Treisman (2008) to capture decision-making decentralization. This indicator is a dummy variable that takes the value one if, under constitution, subnational legislatures have autonomy in certain specified areas not explicitly subject to central laws.[11]

For government quality we rely on the Worldwide Governance Indicators constructed by Kaufmann et al. (1999). These indicators capture six key dimensions of institutional quality: *Voice and accountability*, *Political stability and absence of violence*, *Government effectiveness*, *Regulatory quality*, *Rule of law*, and *Control of corruption*.[12] Nevertheless, these six indicators are characterized by very high bivariate correlations (see Table A6 in the Appendix). Specifically, the lowest correlation is between *Political stability and absence of violence* and *Regulatory quality* ($r = 0.81$), while the highest is between *Rule of law* and *Control of corruption* ($r = 0.99$). Consequently, the six indicators proposed by Kaufmann et al. (1999) appear to be measuring the same broad concept rather than successfully distinguishing between different dimensions of governance (Langbein and Knack, 2010). In order to avoid any potential multicollinearity problem, we follow the strategy adopted by various researchers (e.g. Easterly and Levine, 2003; Seldadyo et al., 2010) and calculate an aggregate index of government quality equal to the average of the six indicators constructed by Kaufmann et al. (1999). This approach also reduces the risk of measurement errors affecting the individual indicators.

We now investigate the relationship between ethnic segregation and the devolution of fiscal and political power from central to subnational governments, and the quality of government. The results are summarized in columns 1-3 of Table 8.[13] Our findings show that ethnic segregation is positively associated with the level of political autonomy of subnational governments, while its effect on the measure of fiscal decen-

[11]It is important to note that, despite their limitations, the measures of fiscal and political decentralization used in our analysis allow us to capture the two dimensions of a well-functioning federal system identified by numerous scholars in the political science literature. For example, Obinger et al. (2005, p. 9) point out that a federal structure comprises "a set of jurisdictional arrangements for allocating policy responsibilities between different levels of government; this refers to both *policy-making* and *policy implementation*." In turn, according to Keman (2000), a federal system includes decentralization with respect to "the right to act" on the one hand, and decentralization with respect to "the right to decide" on the other. The measure of fiscal decentralization employed in our study is related to the implementation of government policies through executing subnational administrations and public good provision ("the right to act"), whereas the indicator of political decentralization reflects the presence of political decision- and law-making power at subnational level.

[12]See Kaufmann et al. (2006) for detailed definitions and sources of each indicator.

[13]The estimates in Table 8 have been obtained by means of 2SLS, despite the fact that the indicator of political decentralization is a dichotomous variable that takes on the values of zero and one. As pointed out by Miguel et al. (2004), 2SLS is typically preferred even in cases in which the dependent variable is dichotomous (Angrist and Krueger, 2001), as strong specification assumptions are required to justify the use of alternative methods, such as those proposed by Rivers and Vuong (1988). In any case, the results are similar when we use an instrumental variable probit model based on a conditional maximum-likelihood estimator.

tralization is not statistically significant.[14] Table 8 also reveals that more ethnically segregated countries have lower government quality, confirming the results obtained by Alesina and Zhuravskaya (2011).

[INSERT TABLE 8 AROUND HERE]

Following the various arguments presented in section 2, we now provide a preliminary test for the potential transmission channels that could explain ultimately the observed relationship between ethnic segregation and spatial inequality. To do this, we include in our baseline model the measures of political decentralization and government quality, which are significantly related to ethnic segregation according to the previous analysis. If these where valid transmission channels, the inclusion of these additional controls should reduce the effect of ethnic segregation on spatial inequality –in terms of coefficient size and/or its statistical significance. Columns 4-7 of Table 8 present the result of the analysis. As can be seen, neither the degree of political decentralization nor government quality exert a statistically significant impact on spatial inequality, conditional on ethnic segregation and the remaining covariates. Nevertheless, the inclusion of these variables influences the relationship between ethnic segregation and spatial inequality. Column 7 shows that, once we control for political decentralization and the quality of government, the coefficient of the index of ethnic segregation still holds positive, but its effect on spatial inequality is statistically significant only at the 10% level. Moreover, in this regression the size of the coefficient of the index of ethnic segregation experiences a 28% decline in magnitude in comparison with the estimates in column 4. While not conclusive, these findings suggest the possibility that political decentralization and the quality of government could be possible transmission channels linking ethnic segregation and spatial inequality, which would be consistent with the arguments discussed in section 2.

In any case, the exploratory nature of this analysis implies that the information provided by Table 8 should be treated with caution. In fact, the two mediating variables considered are highly interdependent, as the degree of decentralization may be associated with the quality of government (Fisman and Gatti, 2002; Enikolopov and Zhuravskaya, 2007). This makes it difficult to isolate the effect of each particular channel. Another important caveat is that political decentralization and the quality of government may themselves be potentially endogenous in this context (Rodríguez-Pose and Ezcurra, 2010; Kyriacou et al., 2015). In order to assess conclusively the relevance of our hypothesized transmission channels, one should exploit an independent exogenous source of variation for each of these variables, a task that we leave open for future investigation.

[14] The F-statistics for the excluded instrument suggest that the index of predicted segregation is weakly correlated with actual segregation in the reduced sample used to estimate the regression in column 1 of Table 8. Nevertheless, we have verified that ethnic segregation is not significantly related to the measure of fiscal decentralization using the weak instrument robust test developed by Finlay and Magnusson (2009).

7 Conclusions

This paper has investigated the link between ethnic segregation and spatial inequality in 71 countries with different levels of economic development. Our results show a positive and statistically significant association between ethnic segregation and the magnitude of within-country regional income disparities. Consequently, more ethnically segregated countries tend on the whole to have higher levels of spatial inequality. The existence of a causal relationship is confirmed by 2SLS regressions using the instrument for segregation constructed by Alesina and Zhuravskaya (2011) and based on the composition of ethnic groups in the home and neighbouring countries. The results still hold when we control for the level of ethnic fractionalization and the stage of economic development. The findings are also robust to the inclusion of other explanatory variables potentially affecting spatial inequality and the geographical distribution of ethnic groups, such as country size, the degree of trade openness, artificial borders, state history, geographical factors, or the redistributive capacity of the public sector. We have also checked that the results are not driven by influential observations or particular groups of countries, and that the findings do not depend on the specific measure used to quantify the levels of spatial inequality and ethnic segregation. Finally, the analysis also suggests the possibility that political decentralization and the quality of government could be possible transmission channels linking ethnic segregation and spatial inequality.

Our research contributes to the existing literature on the determinants of spatial inequality by underlining the role played by ethnic segregation in explaining regional income disparities. The results suggest that the spatial distribution of ethnic groups has important implications for within-country differences in development and therefore should be taken into consideration by policy-makers when designing effective territorial development strategies. This is of particular importance for countries where existing ethnic segregation connected to high levels of spatial inequality make the risk of social unrest and violent armed conflict a real possibility. In particular, the sort of conditions which put internal stability in jeopardy are more likely to take place in low- and middle-income countries, as they generally combine greater levels of both ethnic segregation and spatial inequality.

A number of additional extensions to our work can be derived from the paper. Some relate directly to the enlargement of the number of countries included in the sample, which is particularly important to provide a more complete picture about the nature of the link between ethnic segregation and spatial inequality. Future research will also have to pay special attention to the need to complement and extend our analysis of the various mechanisms which ultimately explain the impact of the geographical distribution of ethnic groups on regional income disparities. Only by pursuing these additional strands, we will be able to attain a fuller grasp of the real implications of ethnic segregation for the evolution of spatial inequality.

References

Abadie, A. and Gardeazabal, J. (2003): The economic costs of conflict: A case study of the Basque Country, *American Economic Review* 93, 113-132.

Acemoglu, D. and Robinson, J.A. (2012): *Why Nations Fail. The Origins of Power, Prosperity, and Poverty.* New York: Crown Business.

Acemoglu, D., Johnson, S. and Robinson, J.A. (2001): The colonial origins of comparative development: An empirical investigation, *American Economic Review* 91, 1369-1401.

Acemoglu, D., Johnson, S. and Robinson, J.A. (2002): Reversal of fortune: Geography and institutions in the making of the modern world income distribution, *Quarterly Journal of Economics* 117, 1231-1294.

Ahlerup, P. and Olsson, O. (2012): The roots of ethnic diversity, *Journal of Economic Growth* 17, 71-102.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S. and Wacziarg, R. (2003): Fractionalization, *Journal of Economic Growth* 8, 155-194.

Alesina, A. and La Ferrara, E. (2005): Ethnic diversity and economic performance, *Journal of Economic Literature* 43, 762-800.

Alesina, A. and Spolaore, E. (1997): On the number and size of nations, *Quarterly Journal of Economics* 112, 1027-1056.

Alesina, A. and Spolaore, E. (2003): *The Size of Nations.* Cambridge (MA): MIT Press.

Alesina, A. and Zhuravskaya, E. (2011): Segregation and the quality of government in a cross section of countries, *American Economic Review* 101, 1872-1911.

Alesina, A., Easterly, W. and Matuszeski, J. (2011): Artificial states, *Journal of the European Economic Association* 9, 246-277.

Alesina, A., Michalopoulos, S. and Papaioannou, E. (2016): Ethnic inequality, *Journal of Political Economy* 124, 428-488.

Angrist, J.D. and Krueger, A.B. (2001): Instrumental variables and the search for identification: from supply and demand to natural experiments, *Journal of Economic Perspectives* 15, 69-85.

Angrist, J.D. and Pischke, J.S. (2009): *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton (NJ): Princeton University Press.

Bakke, K. and Wibbels, E. (2006): Diversity, disparity, and civil conflict in federal states, *World Politics* 59, 1-50.

Baldwin, K. and Huber, J.D. (2010): Economic versus cultural differences: Forms of ethnic diversity and public goods provision, *American Political Science Review* 104, 644-662.

Banerjee, A. and Pande, R. (2007): Parochial politics: Ethnic preferences and politician corruption. Center for Economic and Policy Research Discussion Paper 6381.

Barro, R. and Sala-i-Martin, X. (1995): *Economic Growth.* New York: McGraw Hill.

Bockstette, V., Chanda, A. and Putterman, L. (2002): States and markets: The advantage of an early start, *Journal of Economic Growth* 7, 347-369.

Brancati, D. (2006): Decentralization: Fueling the fire or dampening the flames of ethnic conflict and secessionism?, *International Organization* 60, 651-685.

Cashdan, E. (2001): Ethnic diversity and its environmental determinants: Effects of climate, pathogens, and habitat diversity, *American Anthropologist* 103, 968-991.

Chanda, A. and Putterman, L. (2007): Early starts, reversals and catch-up in the process of economic development, *Scandinavian Journal of Economics* 109, 387-413.

Chandra, K. (2004): *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India.* Cambridge (UK): Cambridge University Press.

Costa, D.L. and Kahn, M.E. (2003): Understanding the decline in American social capital, *Kyklos* 56, 17-46.

Cowell, F. (1995): *Measuring Inequality*, second edition. London: Prentice Hall.

Easterly, W. and Levine, R. (1997): Africa's growth tragedy: Policies and ethnic divisions, *Quarterly Journal of Economics* 112, 1203-1250.

Easterly, W. and Levine, R. (2003): Tropics, germs, and crops: how endowments influence economic development, *Journal of Monetary Economics* 50, 3-39.

Englebert, P., Tarango, S. and Carter, M. (2002): Dismemberment and suffocation. A contribution to the debate on African boundaries, *Comparative Political Studies* 35, 1093-1118.

Enikolopov, R. and Zhuravskaya, E. (2007): Decentralization and political institutions, *Journal of Public Economics* 91, 2261-2290.

Ezcurra, R. and Pascual, P. (2007): Spatial disparities in productivity in Central and Eastern Europe, *Eastern European Economics* 45, 5-32.

Ezcurra, R. and Pascual, P. (2008): Fiscal decentralization and regional disparities: Evidence from several European Union countries, *Environment and Planning A* 40, 1185-1201.

Ezcurra, R. and Rodríguez-Pose, A. (2013): Does economic globalization affect regional inequality? A cross-country analysis, *World Development* 52, 92-103.

Ezcurra, R. and Rodríguez-Pose, A. (2014a): Trade openness and spatial inequality in emerging countries, *Spatial Economic Analysis* 9, 162-182.

Ezcurra, R. and Rodríguez-Pose, A. (2014b): Government quality and spatial inequality: a cross-country analysis, *Environment and Planning A* 2014 46 1732-1753.

Finlay, K. and Magnusson, L.M. (2009): Implementing weak-instrument robust tests for a general class of instrumental-variables models, *The Stata Journal* 9, 398-421.

Fisman, R. and Gatti, R. (2002): Decentralization and corruption: evidence across countries, *Journal of Public Economics* 83, 325-345.

Fujita M. and Thisse J.F. (2002): *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth.* Cambridge (UK): Cambridge University Press.

Gellner, E. (1983): *Nations and Nationalism.* Oxford: Blackwell.

Gennaioli, N., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2013): Human capital and regional development, *Quarterly Journal of Economics* 128, 105-164.

Glaeser, E.L., Laibson, D., Scheinkman, J.A. and Soutter, C.L. (2000): Measuring trust, *Quarterly Journal of Economics* 115, 811-846.

Herbst, J. (2000): *State and Power in Africa.* Princeton (NJ): Princeton University Press.

Hobsbawm, E.J. and Ranger, T. (1983): *The Invention of Tradition.* Cambridge (UK): Cambridge University Press.

Kanbur, R. and Venables, A.J. (eds.) (2005): *Spatial Inequality and Development.* Oxford: Oxford University Press.

Kaufmann, D., Kraay, A. and Zoido-Lobatón, P. (1999): Governance matters. World Bank Policy Research Working Paper 2196, World Bank.

Kaufmann, D., Kraay, A. and Mastruzzi, M. (2006): Governance matters V: Aggregate and individual governance indicators for 1996-2005. World Bank Policy Research Working Paper 4012, World Bank.

Keman, H. (2000): Federalism and policy performance. A conceptual and empirical inquiry. In Wachendorfer-Schmidt, U. (Ed.): *Federalism and Political Performance*, pp. 196-227. London: Routledge.

Krugman, P. (1998): What's new about the new economic geography?, *Oxford Review of Economic Policy* 14, 7-17.

Kyriacou, A.P. and Morral-Palacín, N. (2015): Secessionism and the quality of government: Evidence from a dample of OECD countries, *Territory, Politics, Governance* 3, 187-204.

Kyriacou, A.P. and Roca-Sagalés, O. (2012): The impact of EU structural funds on regional disparities within member states, *Environment and Planning C: Government and Policy* 30, 267-281.

Kyriacou, A.P., Muinelo-Gallo, L. and Roca-Sagalés, O. (2015): Fiscal decentralization and regional disparities: The importance of good governance, *Papers in Regional Science* 94, 89-107.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R. (1997): Trust in large organizations, *American Economic Review* 87, 333-338.

Langbein, L. and Knack, S. (2010): The worldwide governance indicators: Six, one or none?, *Journal of Development Studies* 46, 350-370.

Lessmann, C. (2012): Regional inequality and decentralization - An empirical analysis, *Environment and Planning A* 44, 1363-1388.

Lessmann, C. (2014): Spatial inequality and development - Is there an inverted-U relationship?, *Journal of Development Economics* 106, 35-51.

Matuszeski, J. and Schneider, F. (2006): Patterns of ethnic group segregation and civil conflict, mimeo.

Miguel, E., Satyanath, S. and Sergenti, E. (2004): Economic shocks and civil conflict: an instrumental variables approach, *Journal of Political Economy* 112, 725-753.

Milanovic, B. (2005): *Worlds Apart: Measuring International and Global Inequality.* Princeton (NJ): Princeton University Press.

North, D.C. (1990): *Institutions, Institutional Change and Economic Performance.* New York: Cambridge University Press.

Oates, W. (1972): *Fiscal Federalism.* New York: Harcourt Brace Jovanovich.

Obinger, H.,Leibfried, S. and Castles, F.G. (2005): Introduction: Federalism and the Welfare State. In Obinger, H.,Leibfried, S. and Castles, F.G. (Eds.): *Federalism and the Welfare State: New World and European Experiences*, pp. 1-50. Cambridge (UK): Cambridge University Press.

Østby, G., Nordås, R. and Rød, J.K. (2009): Regional inequalities and civil conflict in Sub-Saharan Africa, *International Studies Quarterly* 53, 301-324.

Petrakos, G., Rodríguez-Pose, A. and Rovolis, A. (2005): Growth, integration and regional inequality in Europe, *Environment and Planning A* 37, 1837-1855.

Portnov, B.A. and Felsentsein, D. (2005): Measures of regional inequality for small countries. In Felsenstein, D. and Portnov, B. (Eds.): *Regional Disparities in Small Countries*, pp. 47-62. Berlin: Springer.

Prud'homme, R. (1995): On the dangers of decentralization. World Bank Policy Research Working Paper 1252, World Bank.

Putnam, R.D. (1993): *Making Democracy Work: Civic Traditions in Modern Italy.* Princeton (NJ): Princeton University Press.

Putterman, L. (2007): State Antiquity Index Dataset, version 3. Brown University.

Qian, Y. and Weingast, B. (1997): Federalism as a commitment to preserving market incentives, *Journal of Economic Perspectives* 11, 83-92.

Ramcharan, R. (2009): Why an economic core: domestic transport costs, *Journal of Economic Geography* 9, 559-581.

Rivers, D. and Vuong, Q.H. (1988): Limited information estimators and exogeneity tests for simultaneous probit models, *Journal of Econometrics* 39, 347-366.

Rodríguez-Pose, A. (2012): Trade and regional inequality, *Economic Geography* 88, 109-136.

Rodríguez-Pose, A. and Ezcurra, R. (2010): Does decentralization matter for regional disparities? A cross-country analysis, *Journal of Economic Geography* 10, 619-644.

Rodríguez-Pose, A. and Gill, N. (2004): Is there a global link between regional disparities and devolution?, *Environment and Planning A* 36, 2097-2117.

Rodríguez-Pose, A. and Gill, N. (2005): On the 'economic dividend' of devolution, *Regional Studies* 39, 405-420.

Rodríguez-Pose, A., Tijmstra, S. and Bwire, A. (2009): Fiscal decentralisation, efficiency, and growth, *Environment and Planning A* 41, 2041-2062.

Schneider, A. (2003): Decentralization: conceptualization and measurement, *Studies in Comparative International Development* 38, 32-56.

Seldadyo, J., Elhorst, P. and De Haan, J. (2010): Geography and governance: Does space matter?, *Papers in Regional Science* 89, 625-640.

Shankar, R. and Shah, A. (2003): Bridging the economic divide within countries: A scorecard on the performance of regional policies in reducing regional income disparities, *World Development* 31, 1421-1441.

Sorens, J. (2005): The cross-sectional determinants of secessionism in advanced democracies, *Comparative Political Studies* 38, 304-326.

Staiger, D. and Stock, J.H. (1997): Instrumental variables regression with weak instruments, *Econometrica* 65, 557-586.

Stock, J.H., Wright, J.H. and Yogo, M. (2002): A survey of weak instruments and weak identification in generalized method of moments, *Journal of Business and Economic Statistics* 20, 518-529.

Thisse, J.F. (2000): Agglomeration and regional imbalance: Why? And is it bad?, *Papers of the European Investment Bank* 5, 47-67.

Tiebout, C.M. (1956): A pure theory of local expenditures, *Journal of Political Economy* 64, 416-424.

Tilly, C. (1992): *Coercion, Capital and European States, AD 990-1992*. Oxford: Blackwell.

Treisman, D. (2008): Decentralization dataset. Available at: http://www.sscnet.ucla.edu/polisci/faculty/treisman/.

Uslaner, E.M. (2008): Trust, diversity and segregation. Unpublished manuscript.

Weingast, B. (1995): The economic role of political institutions: Market-preserving federalism and economic growth, *Journal of Law, Economics and Organization* 11, 1-31.

Williamson, J. G. (1965): Regional inequality and the process of national development: A description of patterns, *Economic Development and Cultural Change* 13, 3-45.

Tables

Table 1: The most and the least segregated countries.

Most segregated countries		Least segregated countries	
Country	$\hat{S}$	Country	$\hat{S}$
Zimbabwe	0.394	Germany	0.001
Guatemala	0.384	Sweden	0.001
Turkey	0.357	Netherlands	0.001
Ecuador	0.322	Korea	0.002
Pakistan	0.316	Japan	0.002
Colombia	0.280	Greece	0.002
Morocco	0.253	Paraguay	0.002
South Africa	0.247	Slovenia	0.003
Spain	0.244	Hungary	0.003
Kenya	0.235	Ireland	0.003

Note: The measure of ethnic segregation was drawn from Alesina and Zhuravskaya (2011).

Table 2: The most and the least unequal countries.

Most unequal countries		Least unequal countries	
Country	c	Country	c
Ecuador	1.112	Australia	0.076
Russia	0.931	Malawi	0.121
Kenya	0.910	Netherlands	0.137
Indonesia	0.891	Greece	0.156
Guatemala	0.700	Senegal	0.163
Kazakhstan	0.688	United States	0.167
Philippines	0.622	Canada	0.170
Mexico	0.602	Finland	0.171
Latvia	0.593	Armenia	0.186
China	0.586	New Zealand	0.188

Note: See Table A1 in the Appendix for further details on the sources of regional data on GDP and population used to calculate c.

Table 3: Spatial inequality and ethnic segregation: OLS regressions.

	(1)	(2)	(3)	(4)	(5)	(6)
Ethnic segregation	1.952***	1.672***	1.727***	1.467***	1.790***	1.532***
	(0.550)	(0.578)	(0.521)	(0.457)	(0.520)	(0.457)
Average size of regions			-0.223***	-0.256***	-0.226***	-0.260***
			(0.065)	(0.070)	(0.070)	(0.073)
Ethnic fractionalization				0.439**		0.449**
				(0.179)		(0.184)
GDP per capita					0.119	0.126*
					(0.078)	(0.068)
Urban population					-0.006	-0.006*
					(0.004)	(0.003)
Trade openness			0.509***	0.528***	0.529***	0.549***
			(0.145)	(0.144)	(0.137)	(0.134)
Area			0.334***	0.359***	0.345***	0.371***
			(0.068)	(0.071)	(0.069)	(0.071)
Transition			0.391***	0.362***	0.446***	0.420***
			(0.101)	(0.099)	(0.110)	(0.102)
Constant	-1.281***	-1.900***	-4.310***	-4.426***	-5.058***	-5.231***
	(0.069)	(0.146)	(0.587)	(0.543)	(0.839)	(0.739)
Continent dummies	No	Yes	Yes	Yes	Yes	Yes
R-squared	0.162	0.406	0.671	0.701	0.686	0.719
Observations	71	71	71	71	71	71

Notes: The dependent variable is in all cases the log of c. Robust standard errors in parentheses. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 4: First stage regressions: Segregation and predicted segregation.

	(1)	(2)	(3)	(4)	(5)	(6)
Predicted ethnic segreg.	0.623***	0.578***	0.566***	0.543***	0.582***	0.563***
	(0.153)	(0.141)	(0.130)	(0.135)	(0.146)	(0.146)
Average size of regions			-0.018	-0.022	-0.019	-0.023
			(0.016)	(0.014)	(0.016)	(0.014)
Ethnic fractionalization				0.074		0.072
				(0.045)		(0.046)
GDP per capita					0.012	0.013
					(0.023)	(0.022)
Urban population					-0.001	-0.001
					(0.001)	(0.001)
Trade openness			0.002	0.006	0.008	0.011
			(0.038)	(0.034)	(0.039)	(0.036)
Area			0.018	0.022	0.021	0.025
			(0.016)	(0.015)	(0.016)	(0.015)
Transition			-0.049**	-0.051***	-0.046*	-0.048**
			(0.019)	(0.018)	(0.024)	(0.023)
Constant	0.051***	0.172***	0.111	0.086	0.059	0.025
	(0.011)	(0.020)	(0.131)	(0.129)	(0.193)	(0.194)
Continent dummies	No	Yes	Yes	Yes	Yes	Yes
F-stat. (hom.)	41.03***	38.39***	35.71***	33.22***	30.42***	28.93***
F-stat. (het.)	16.52***	16.87***	18.82***	16.09***	15.99***	14.89***
Partial R-squared	0.373	0.379	0.377	0.364	0.348	0.341
R-squared	0.373	0.499	0.535	0.556	0.557	0.577
Observations	71	71	71	71	71	71

Notes: The dependent variable is in all cases the index of ethnic segregation $\hat{S}$. Robust standard errors in parentheses. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 5: Spatial inequality and ethnic segregation: Second stage regressions.

	(1)	(2)	(3)	(4)	(5)	(6)
Predicted ethnic seg.	2.036**	1.870**	1.556**	1.309**	2.001***	1.808***
	(0.823)	(0.800)	(0.684)	(0.561)	(0.720)	(0.604)
Average size of regions			-0.227***	-0.261***	-0.221***	-0.251***
			(0.061)	(0.065)	(0.066)	(0.068)
Ethnic fractionalization				0.456***		0.422**
				(0.173)		(0.175)
GDP per capita					0.128*	0.138**
					(0.072)	(0.061)
Urban population					-0.006*	-0.006**
					(0.003)	(0.003)
Trade openness			0.506***	0.526***	0.533***	0.553***
			(0.135)	(0.132)	(0.124)	(0.120)
Area			0.339***	0.364***	0.339***	0.362***
			(0.064)	(0.067)	(0.066)	(0.067)
Transition			0.385***	0.355***	0.460***	0.439***
			(0.093)	(0.091)	(0.105)	(0.096)
Constant	-1.288***	-1.950***	-4.281***	-4.406***	-5.166***	-5.353***
	(0.094)	(0.203)	(0.535)	(0.485)	(0.782)	(0.659)
Continent dummies	No	Yes	Yes	Yes	Yes	Yes
R-squared	0.161	0.405	0.670	0.701	0.685	0.716
Observations	71	71	71	71	71	71

Notes: The dependent variable is in all cases the log of c. Robust standard errors in parentheses. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 6: Robustness analysis: Impact of different groups of countries.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnic segregation	1.449**	1.798***	1.785**	2.117***	2.148***	1.895***	1.762***
	(0.585)	(0.608)	(0.839)	(0.806)	(0.618)	(0.619)	(0.605)
Omitted countries	Latin America	North America	Europe	Sub-Sah. Africa	Asia	Pacific	MENA
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.737	0.703	0.744	0.740	0.673	0.709	0.713
F-stat. (hom.)	31.27***	28.35***	18.17***	15.78***	23.47***	28.24***	27.88***
F-stat. (het.)	12.25***	14.95***	20.48***	8.93***	13.34***	14.94***	14.90***
Observations	58	69	44	59	57	69	69

Notes: Second stage regressions. The dependent variable is in all cases the log of c. Robust standard errors in parentheses. All the regressions include continent dummies and the full set of control variables described in section 4.1. 'F-stat. (hom.)' and 'F-stat. (het.)' report the F-statistics for the excluded instrument from the first stage under the assumptions of homoskedasticity and heteroskedasticity respectively. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 7: Robustness analysis: Additional controls.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Ethnic segregation	2.072***	2.273***	1.863***	1.821***	1.938***	2.032***	2.157***	1.690***	1.728***	1.482**	1.736***
	(0.618)	(0.692)	(0.611)	(0.604)	(0.549)	(0.735)	(0.618)	(0.621)	(0.642)	(0.603)	(0.559)
Squiggliness	-2.175										
	(2.300)										
Partitioned		-0.003									
		(0.002)									
State antiquity			-0.135								
			(0.094)								
Former colony				0.024							
				(0.183)							
Spanish colony					0.257*						
					(0.150)						
British colony					-0.266*						
					(0.158)						
French colony					-0.261						
					(0.207)						
Portuguese colony					0.324						
					(0.218)						
Number of ethnic groups						0.037					
						(0.052)					
Ethnic segr. neighb. countries							-0.031				
							(0.537)				
Ethnic frac. neigh. countries							-0.116				
							(0.211)				
Ruggedness								0.040			
								(0.034)			
Mountains									0.140		
									(0.166)		
Rivers										-0.028***	
										(0.008)	
Non-contiguous territory											-0.043
											(0.110)
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.713	0.712	0.740	0.716	0.739	0.713	0.709	0.723	0.720	0.736	0.718
F-stat. (hom.)	29.54***	22.57***	30.63***	28.18***	32.72***	19.55***	29.10***	27.22***	26.935***	25.62***	30.48***
F-stat. (het.)	14.83***	14.97***	17.15***	13.74***	15.73***	11.96***	13.11***	13.99***	13.28***	13.40***	15.17***
Observations	65	60	69	71	71	71	65	71	71	71	71

Notes: Second stage regressions. The dependent variable is in all cases the log of c. Robust standard errors in parentheses. All the regressions include continent dummies and the full set of control variables described in section 4.1. 'F-stat. (hom.)' and 'F-stat. (het.)' report the F-statistics for the excluded instrument from the first stage under the assumptions of homoskedasticity and heteroskedasticity respectively. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 7: Robustness analysis: Additional controls. (*Continuation.*)

	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
Ethnic segregation	1.812***	1.827***	1.831***	1.828***	1.836***	1.722***	1.796***	1.683***	1.820***	1.981***	1.818***
	(0.594)	(0.602)	(0.592)	(0.593)	(0.605)	(0.661)	(0.588)	(0.552)	(0.649)	(0.670)	(0.604)
Landlocked	-0.019										
	(0.079)										
Island		0.218									
		(0.161)									
Distance to the coast			-0.048								
			(0.106)								
Latitude				-0.002							
				(0.004)							
Number of neighb. countries					-0.014						
					(0.015)						
Ethnic inequality (GREG)						-0.173					
						(0.200)					
Ethnic inequality (Ethnol.)							-0.261				
							(0.198)				
Ethnic tensions								-0.016			
								(0.036)			
Internal conflict									-0.014		
									(0.119)		
Public investment										-0.002	
										(0.016)	
Public consumption											0.209
											(0.624)
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.717	0.720	0.717	0.717	0.719	0.720	0.723	0.739	0.716	0.717	0.717
F-stat. (hom.)	30.06***	28.37***	28.91***	29.61***	28.50***	26.45***	30.75***	27.67***	27.43***	24.38***	28.18***
F-stat. (het.)	13.29***	14.69***	13.74***	13.81***	14.44***	13.94***	16.17***	15.77***	14.41***	14.27***	14.42***
Observations	71	71	71	71	71	71	71	68	71	68	71

Notes: Second stage regressions. The dependent variable is in all cases the log of c. Robust standard errors in parentheses. All the regressions include continent dummies and the full set of control variables described in section 4.1. 'F-stat. (hom.)' and 'F-stat. (het.)' report the F-statistics for the excluded instrument from the first stage under the assumptions of homoskedasticity and heteroskedasticity respectively. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 8: Spatial inequality and ethnic segregation: Transmission channels.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	Fiscal decentral.	Political autonomy	Government quality	Spatial inequality	Spatial inequality	Spatial inequality	Spatial inequality
Ethnic segregation	0.876	2.060***	-2.196***	1.589***	1.373**	1.413**	1.136*
	(0.636)	(0.771)	(0.675)	(0.563)	(0.543)	(0.630)	(0.632)
Political autonomy					0.105		0.114
					(0.095)		(0.092)
Government quality						-0.080	-0.099
						(0.111)	(0.119)
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.474	0.174	0.891	0.740	0.742	0.740	0.741
F-stat. (hom.)	5.72**	33.28***	28.93***	33.28***	29.47***	26.22***	22.72***
F-stat. (het.)	9.47***	14.21***	14.88***	14.21***	11.23***	10.99***	8.57***
Observations	46	63	71	63	63	63	63

Notes: Second stage regressions. Robust standard errors in parentheses. All the regressions include continent dummies and the full set of control variables described in section 4.1. 'F-stat. (hom.)' and 'F-stat. (het.)' report the F-statistics for the excluded instrument from the first stage under the assumptions of homoskedasticity and heteroskedasticity respectively. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

APPENDIX

Definitions and sources of control variables

Average size of regions: Log of the average area of the country's regions expressed in square kilometres. Source: Own elaboration.

Ethnic fractionalization: Index of ethnic fractionalization based on aggregated regional data. The index captures the probability that two individuals randomly drawn from the population belong to different ethnic groups. Source: Alesina and Zhuravskaya (2011).

Linguistic fractionalization: Index of linguistic fractionalization based on aggregated regional data. The index captures the probability that two individuals randomly drawn from the population belong to different linguistic groups. Source: Alesina and Zhuravskaya (2011).

Religious fractionalization: Index of religious fractionalization based on aggregated regional data. The index captures the probability that two individuals randomly drawn from the population belong to different religious groups. Source: Alesina and Zhuravskaya (2011).

GDP per capita: Log of GDP per capita based on purchasing power parity and expressed in constant 2005 international dollars. Average of the years 1995-2005. Source: Penn World Table 7.1. (Heston et al., 2012).

Urban population: Fraction of the total population living in urban areas. Average of the years 1995-2005. Source: World Development Indicators.

Trade openness: Sum of exports and imports of goods and services measured as a share of GDP. Average of the years 1995-2005. Source: World Development Indicators.

Area: Log of country's total area expressed in square kilometres. Source: World Development Indicators.

Transition: Dummy variable that takes the value one if the country is a transition economy, zero otherwise. Source: Alesina and Zhuravskaya (2011).

Squiggliness: Log of fractal dimension of country's non-coastline border. Source: Alesina et al. (2011).

Partitioned: Fraction of country's population belonging to ethnic groups partitioned by a border. Source: Alesina et al. (2011).

State antiquity: Log of discounted sum of years of ancient statehood, 1 AD to 1950 AD. Source: Putterman (2007).

Former colony: Dummy variable that takes the value one if the country is a former European colony, zero otherwise. Source: Nunn and Puga (2012).

British colony: Dummy variable that takes the value one if the country is a former British colony, zero otherwise. Source: Nunn and Puga (2012).

French colony: Dummy variable that takes the value one if the country is a former French colony, zero otherwise. Source: Nunn and Puga (2012).

Spanish colony: Dummy variable that takes the value one if the country is a former Spanish colony, zero otherwise. Source: Nunn and Puga (2012).

Portuguese colony: Dummy variable that takes the value one if the country is a former Portuguese colony, zero otherwise. Source: Nunn and Puga (2012).

Number of ethnic groups: Log of number of ethnic groups in each country. Source: Alesina and Zhuravskaya (2011).

Ethnic segregation in neighbouring countries: Population-weighted average of the degree of ethnic segregation in neighbouring countries. Ethnic segregation in each country is calculated according to expression (2). Source: Alesina and Zhuravskaya (2011).

Ethnic fractionalization in neighbouring countries: Population-weighted average of the degree of ethnic fractionalization in neighbouring countries. Ethnic fractionalization in each country captures the probability that two individuals randomly drawn from the population belong to different ethnic groups. Source: Alesina and Zhuravskaya (2011).

Ruggedness: Index of terrain ruggedness. Source: Nunn and Puga (2012).

Mountains: Measure of the extent to which a country's surface is covered by mountains. Source: Alesina and Zhuravskaya (2011).

Rivers: Measure of the extent to which a country is covered by rivers or other inland bodies of water. Source: Alesina and Zhuravskaya (2011).

Non-contiguous state: Dummy variable that takes the value one for those countries with territory holding at least 10,000 people and separated from the land area containing the capital city either by land or by 100 kilometers of water, zero otherwise. Source: Fearon and Laitin (2003).

Landlocked: Dummy variable that takes the value one if the country is entirely enclosed by land, or whose only coastlines lie on closed seas, zero otherwise. Source: Global Development Network Growth Database (World Bank).

Island: Dummy variable that takes the value one if the country is surrounded by water and has no bordering countries, and zero otherwise. Source: Alesina and Zhuravskaya (2011).

Distance to coast: Average distance to nearest ice-free coast (1000 km). Source: Nunn and Puga (2012).

Latitude: Absolute value of the latitude of the country. Source: Own elaboration with data drawn from Nunn and Puga (2012).

Number of neighbouring countries: Number of neighbouring countries. Source: CIA, The World Factbook.

Ethnic inequality (GREG): Gini index that reflects the differences in mean income –as captured by luminosity per capita at the ethnic homeland– across ethnic groups. The location of the various ethnic groups is identified using the Geo-Referencing of

Ethnic Groups (GREG), which is the digitized version of the Soviet Atlas Narodov Mira. Source: Alesina et al. (2016).

Ethnic inequality (Ethnologue): Gini index that reflects the differences in mean income –as captured by luminosity per capita at the ethnic homeland– across ethnic groups. The location of the various ethnic groups is identified using the Ethnologue. Source: Alesina et al. (2016).

Ethnic tensions: Index of the degree of tension within a country attributable to racial, nationality, or language divisions. Lower values indicate more tensions. Average of the years 1995-2005. Source: International Country Risk Guide (ICRG).

Civil conflict: Dummy variable that takes the value one if the country has experienced a civil armed conflict between 1995 and 2005, and zero otherwise. A country is recorded as having experienced a civil armed conflict in a given year if a threshold of 25 or more battle-related deaths has been met. Source: UCDP/PRIO.

Public investment: General government investment (gross fixed capital formation), expressed as a share of GDP. Average of the years 1995-2005. Source: IMF Investment and Capital Stock Dataset (International Monetary Fund, 2015).

Public consumption: General government final consumption expenditure expressed as a share of GDP. Average of the years 1995-2005. Source: World Development Indicators.

Fiscal decentralization: Measure of fiscal decentralization obtained using factor analysis of data from 1996 on six indicators: subnational expenditures as a percentage of total expenditures, subnational revenues as a percentage of total revenues, relative importance of taxes as a percentage of subnational revenues, relative importance of transfers as a percentage of subnational revenues, existence of municipal elections, and existence of state or provincial elections. Source: Schneider (2003).

Political autonomy: Dummy variable that takes the value one if, under constitution, subnational legislatures have autonomy in certain specified areas not explicitly subject to central laws, zero otherwise. Source: Treisman (2008).

Government quality: Average of the six indicators of the quality of government proposed by Kaufmann et al. (1999). The six indicators are: Voice and accountability, Political stability and absence of violence, Government effectiveness, Regulatory quality, Rule of law, and Control of corruption. Average of the period 1996-2005. Source: Alesina and Zhuravskaya (2011).

References

Alesina, A. and Zhuravskaya, E. (2011): Segregation and the quality of government in a cross section of countries, *American Economic Review* 101, 1872-1911.

Alesina, A., Easterly, W. and Matuszeski, J. (2011): Artificial states, *Journal of the European Economic Association* 9, 246-277.

Alesina, A., Michalopoulos, S. and Papaioannou, E. (2016): Ethnic inequality, *Journal* *of Political Economy* 124, 428-488.

Ezcurra, R. and Rodríguez-Pose, A. (2013): Does economic globalization affect regional inequality? A cross-country analysis, *World Development* 52, 92-103.

Fearon, J. and Laitin, D. (2003): Ethnicity, insurgency, and civil war, *American Political Science Review* 97, 75-90.

Gennaioli, N., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2013): Human capital and regional development, *Quarterly Journal of Economics* 128, 105-164.

Heston, A. Summers, R. and Aten, B. (2012): Penn World Table Version 7.1. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

International Monetary Fund (2015): Making Public Investment more Efficient. IMF Staff Report. Washington D.C., International Monetary Fund.

Kaufmann, D., Kraay, A. and Zoido-Lobatón, P. (1999): Governance matters. World Bank Policy Research Working Paper 2196, World Bank.

Nunn, N. and Puga, D. (2012): Ruggedness: The blessing of bad geography in Africa, *Review of Economics and Statistics* 94, 20-36.

Putterman, L. (2007): State Antiquity Index Dataset, version 3. Brown University.

Treisman, D. (2008): Decentralization dataset. Available at: http://www.sscnet.ucla.edu/polisci/faculty/treisman/.

Figures and Tables

Figure A1: Partial regression plot: Spatial inequality and ethnic segregation.

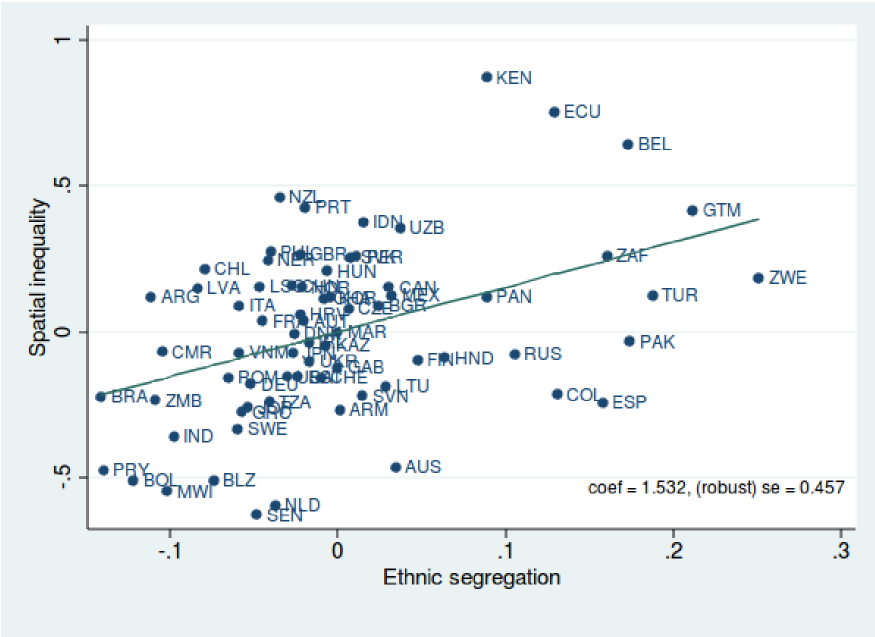

Figure A2: Partial regression plot: Segregation and predicted segregation.

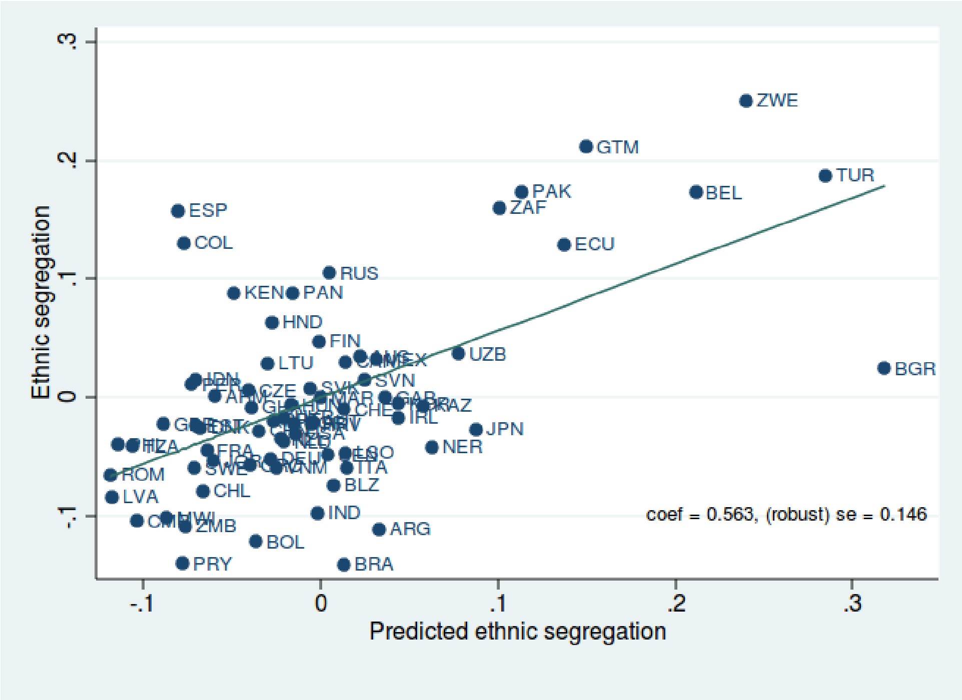

Table A1: Sources of regional data on GDP and population.

Country	Source	Country	Source
Argentina	Ezcurra and Rodríguez-Pose (2013)	Korea	Ezcurra and Rodríguez-Pose (2013)
Armenia	Gennaioli et al. (2013)	Latvia	Gennaioli et al. (2013)
Australia	Ezcurra and Rodríguez-Pose (2013)	Lesotho	Gennaioli et al. (2013)
Austria	Ezcurra and Rodríguez-Pose (2013)	Lithuania	Ezcurra and Rodríguez-Pose (2013)
Belgium	Ezcurra and Rodríguez-Pose (2013)	Malawi	Gennaioli et al. (2013)
Belize	Gennaioli et al. (2013)	Mexico	Ezcurra and Rodríguez-Pose (2013)
Bolivia	Ezcurra and Rodríguez-Pose (2013)	Morocco	Gennaioli et al. (2013)
Brazil	Ezcurra and Rodríguez-Pose (2013)	Netherlands	Ezcurra and Rodríguez-Pose (2013)
Bulgaria	Ezcurra and Rodríguez-Pose (2013)	New Zealand	Ezcurra and Rodríguez-Pose (2013)
Cameroon	Gennaioli et al. (2013)	Niger	Gennaioli et al. (2013)
Canada	Ezcurra and Rodríguez-Pose (2013)	Norway	Ezcurra and Rodríguez-Pose (2013)
Chile	Ezcurra and Rodríguez-Pose (2013)	Pakistan	Gennaioli et al. (2013)
China	Ezcurra and Rodríguez-Pose (2013)	Panama	Gennaioli et al. (2013)
Colombia	Ezcurra and Rodríguez-Pose (2013)	Paraguay	Gennaioli et al. (2013)
Croatia	Eurostat	Peru	Ezcurra and Rodríguez-Pose (2013)
Czech Republic	Ezcurra and Rodríguez-Pose (2013)	Philippines	Ezcurra and Rodríguez-Pose (2013)
Denmark	Ezcurra and Rodríguez-Pose (2013)	Portugal	Ezcurra and Rodríguez-Pose (2013)
Ecuador	Ezcurra and Rodríguez-Pose (2013)	Romania	Ezcurra and Rodríguez-Pose (2013)
Estonia	Gennaioli et al. (2013)	Russia	Gennaioli et al. (2013)
Finland	Ezcurra and Rodríguez-Pose (2013)	Senegal	Gennaioli et al. (2013)
France	Ezcurra and Rodríguez-Pose (2013)	Slovakia	Ezcurra and Rodríguez-Pose (2013)
Gabon	Gennaioli et al. (2013)	Slovenia	Ezcurra and Rodríguez-Pose (2013)
Germany	Ezcurra and Rodríguez-Pose (2013)	South Africa	Ezcurra and Rodríguez-Pose (2013)
Ghana	Gennaioli et al. (2013)	Spain	Ezcurra and Rodríguez-Pose (2013)
Greece	Ezcurra and Rodríguez-Pose (2013)	Sweden	Ezcurra and Rodríguez-Pose (2013)
Guatemala	National statistics	Switzerland	Ezcurra and Rodríguez-Pose (2013)
Honduras	Gennaioli et al. (2013)	Tanzania	Gennaioli et al. (2013)
Hungary	Ezcurra and Rodríguez-Pose (2013)	Turkey	Ezcurra and Rodríguez-Pose (2013)
India	Ezcurra and Rodríguez-Pose (2013)	Ukraine	Gennaioli et al. (2013)
Indonesia	Ezcurra and Rodríguez-Pose (2013)	United Kingdom	Ezcurra and Rodríguez-Pose (2013)
Ireland	Ezcurra and Rodríguez-Pose (2013)	United States	Ezcurra and Rodríguez-Pose (2013)
Italy	Ezcurra and Rodríguez-Pose (2013)	Uzbekistan	Gennaioli et al. (2013)
Japan	Ezcurra and Rodríguez-Pose (2013)	Vietnam	Gennaioli et al. (2013)
Jordan	Gennaioli et al. (2013)	Zambia	Gennaioli et al. (2013)
Kazakhstan	National statistics	Zimbabwe	Gennaioli et al. (2013)
Kenya	Gennaioli et al. (2013)		

Table A2: Spatial inequality in various groups of countries.

	Two groups		Three groups		
	Low segregation	High segregation	Low segregation	Middle segregation	High segregation
Spatial inequality	0.307	0.445	0.275	0.343	0.536
Countries	36	35	18	35	18
Test of equality of means (p-value)	9.35 (0.003)			10.70 (0.000)	

Notes: The classifications are based on the median (classification into two groups) and the first and third quartiles (classification into three groups) of the distribution of $\hat{S}$. Spatial inequality is calculated for each country according to expression (4).

Table A3: Linguistic segregation, religious segregation, and spatial inequality.

	(1)	(2)	(3)	(4)
Method	0LS	OLS	2SLS	2SLS
Linguistic segregation	0.774*		0.878	
	(0.459)		(0.610)	
Linguistic fractionalization	0.266		0.254	
	(0.228)		(0.204)	
Religious segregation		1.132		1.559
		(1.110)		(1.144)
Religious fractionalization		0.012		-0.013
		(0.348)		(0.318)
Average size of regions	-0.273***	-0.234**	-0.269***	-0.218***
	(0.080)	(0.100)	(0.077)	(0.077)
GDP per capita	0.072	0.062	0.077	0.064
	(0.089)	(0.113)	(0.080)	(0.097)
Urban population	-0.005	-0.006	-0.005	-0.006
	(0.005)	(0.005)	(0.004)	(0.005)
Trade openness	0.572***	0.400*	0.579***	0.381**
	(0.174)	(0.222)	(0.160)	(0.188)
Area	0.376***	0.368***	0.373***	0.345***
	(0.079)	(0.088)	(0.075)	(0.074)
Transition	0.284**	0.263*	0.290**	0.257*
	(0.122)	(0.154)	(0.113)	(0.131)
Constant	-4.538***	-3.900***	-4.589***	-4.941***
	(0.979)	(0.996)	(0.884)	(1.072)
Continent dummies	Yes	Yes	Yes	Yes
R-squared	0.631	0.648	0.631	0.646
F-stat. (hom.)			14.95***	17.74***
F-stat. (het.)			12.44***	4.77**
Observations	64	49	64	49

Notes: The dependent variable is in all cases the log of c. Robust standard errors in parentheses. 'F-stat. (hom.)' and 'F-stat. (het.)' report the F-statistics for the excluded instrument from the first stage under the assumptions of homoskedasticity and heteroskedasticity respectively. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table A4: Spatial inequality and ethnic segregation: Quantile regressions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Quantile	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Ethnic segregation	1.406***	1.074**	1.631***	1.187**	1.222**	1.647***	1.945***	2.311***	2.867***
	(0.392)	(0.515)	(0.539)	(0.581)	(0.499)	(0.508)	(0.506)	(0.571)	(0.445)
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.557	0.515	0.501	0.49	0.49	0.501	0.506	0.51	0.554
Observations	71	71	71	71	71	71	71	71	71

Notes: The dependent variable is in all cases the log of c. All the regressions include continent dummies and the full set of control variables described in section 4.1. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table A5: Robustness analysis: Alternative measures of spatial inequality and ethnic segregation.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	c	$T(1)$	s	c	$T(1)$	s
Ethnic segregation ($\hat{S}$)	1.808***	3.336***	1.481***			
	(0.604)	(1.003)	(0.476)			
Ethnic segregation ($\overline{S}$)				1.846**	3.428***	1.550**
				(0.779)	(1.330)	(0.623)
Continent dummies	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.716	0.732	0.735	0.689	0.700	0.700
F-stat. (hom.)	28.93***	28.93***	28.93***	12.90***	12.90***	12.90***
F-stat. (het.)	14.89***	14.89***	14.89***	8.85***	8.85***	8.85***
Observations	71	71	71	71	71	71

Notes: The dependent variables are expressed in logs. Second stage regressions. Robust standard errors in parentheses. 'F-stat. (hom.)' and 'F-stat. (het.)' report the F-statistics for the excluded instrument from the first stage under the assumptions of homoskedasticity and heteroskedasticity respectively. All the regressions include continent dummies and the full set of control variables described in section 4.1. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table A6: Correlation coefficients between the various dimensions of government quality.

	Voice and accountability	Political stability	Government effectiveness	Regulatory quality	Rule of law	Control of corruption
Voice and accountability	1.000					
Political stability	0.813	1.000				
Government effectiveness	0.868	0.813	1.000			
Regulatory quality	0.915	0.812	0.933	1.000		
Rule of law	0.887	0.855	0.981	0.924	1.000	
Control of corruption	0.870	0.827	0.985	0.916	0.986	1.000

Note: The value of each indicator is the mean over the period 1996-2005. All the correlation coefficients are statistically significant at the 1% level.